

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan hasil penelitian implementasi *random forest* dengan *parallel computing* dalam R, didapat beberapa kesimpulan berikut ini:

1. Perancangan *random forest* untuk *parallel computing* dilakukan dengan cara membagi sejumlah *tree* ke beberapa *processor* kemudian dilakukan prediksi..
2. Metode *random forest* dengan *parallel computing* diimplementasikan dalam bahasa pemrograman R dengan menggunakan *package foreach* dan *doParallel*.
3. Implementasi metode *random forest* dengan *parallel computing* diuji menggunakan *dataset* bunga Iris, kualitas *wine* dan data diabetes wanita Pima Indian menggunakan beberapa skenario pengujian. Skenario pengujian dilakukan dengan memodifikasi jumlah *processor* yang digunakan, jumlah *tree* dan jumlah fiturnya lalu dihitung dan dianalisa waktu komputasinya. Berdasarkan hasil prediksi *spesies* bunga Iris dengan metode *random forest* konvensional didapat waktu tersingkat 1.304862 detik yaitu pada saat numTree 20 dan numFeature 2 dan *error* yang dihasilkan 10,00%. Sementara waktu tersingkat yang didapat saat menggunakan metode *parallel random forest* adalah 1.327446 detik pada saat numProcessor 3, numTree 20 dan numFeature 2 serta *error* yang dihasilkan pun sama yaitu 10,00%. Pada kasus prediksi kualitas *wine* waktu tersingkat yaitu 5.147817 detik dengan menggunakan metode *random forest* konvensional saat numTree 20 dan numFeature 2, *error* yang dihasilkan 0,00%. Waktu tersingkat saat menggunakan metode *parallel random forest* adalah 2.570818 detik saat numProcessor 4, numTree 20 dan numFeature 2. *Error* yang dihasilkan juga sama 0,00%. Metode *random forest* konvensional saat memprediksi diabetes wanita Pima Indian menghasilkan waktu tersingkat yaitu 31.39034 detik pada

saat numTree 20 dan numFeature 2, *error* yang dihasilkan yaitu 37.6224%. Sementara saat menggunakan *parallel random forest*, waktu tersingkatnya adalah 9.927734 detik saat numProcessor 4, numTree 20 dan numFeature 2. *Error* yang dihasilkan pun lebih kecil dari *random forest* konvensional, yaitu 29.078%. Berdasarkan data tersebut dapat disimpulkan bahwa waktu komputasi *random forest* dengan *parallel computing* jauh lebih cepat dibandingkan dengan *random forest* konvensional dan hasil prediksi tidak mempunyai perbedaan yang signifikan.

5.2 Saran

Penelitian ini masih memiliki banyak kekurangan yang bisa diperbaiki oleh peneliti lain yang akan mengembangkan penelitian serupa. Oleh karena itu beberapa hal ini dapat menjadi rujukan peneliti selanjutnya agar hasil penelitian menjadi lebih baik, diantaranya:

1. Penelitian ini masih menggunakan satu komputer dengan beberapa *processor*. Untuk memproses data yang lebih banyak dan pemrosesan yang lebih cepat, bisa digunakan beberapa komputer dengan metode *parallel computing*.
2. Sistem ini masih belum memiliki tampilan yang memadai atau masih dalam bentuk *console*, diharapkan untuk penelitian berikutnya tampilannya bisa menggunakan *graphical user interface* agar lebih baik.